Supplementary Appendix

This appendix has been provided by the authors to give readers additional information about their work.

Supplement to: Collier JJ, Guissart C, Oláhová M, et al. Developmental consequences of defective ATG7-mediated autophagy in humans. N Engl J Med 2021;384:2406-17. DOI: 10.1056/NEJMoa1915722

Developmental Consequences of Defective ATG7-mediated Autophagy in Humans

Jack J Collier, Ph.D., Claire Guissart, Pharm.D., Monika Oláhová, Ph.D., Souphatta Sasorith, Ph.D., Florence Piron-Prunier, M.Sc., Fumi Suomi, Ph.D., David Zhang, M.Sc., Nuria Martinez-Lopez, Ph.D., Nicolas Leboucq, M.D., Angela Bahr, Ph.D., Silvia Azzarello-Burri, M.D., Selina Reich, M.Sc., Ludger Schöls, M.D., Tuomo M Polvikoski, Ph.D., Pierre Meyer, M.D., Lise Larrieu, M.Sc., Andrew M Schaefer, M.R.C.P., Hessa S Alsaif, B.Sc., Suad Alyamani, M.D., Stephan Zuchner, Ph.D., Inês A Barbosa, Ph.D., Charu Deshpande, F.R.C.P., Angela Pyle, Ph.D., Anita Rauch, M.D., Matthis Synofzik, M.D., Fowzan S Alkuraya, M.D., François Rivier, M.D., Mina Ryten, Ph.D., Robert McFarland, Ph.D., Agnès Delahodde, Ph.D., Thomas G McWilliams, Ph.D., Michel Koenig, M.D. and Robert W Taylor, Ph.D.

The author's affiliations are as follows: Wellcome Centre for Mitochondrial Research, (J.J.C., M.O., N.M-L, A.M.S, A.P., R.M., R.W.T.), Translational and Clinical Research Institute (J.J.C, M.O., T.M.P., A.M.S., A.P., R.M., R.W.T.), NHS Highly Specialised Service for Rare Mitochondrial Disorders of Adults and Children (A.M.S., R.M., R.W.T.) – all at Newcastle University, Newcastle upon Tyne, UK; Institut Universitaire de Recherche Clinique, and Laboratoire de Génétique Moléculaire, University of Montpellier and Centre Hospitalier Universitaire (CHU) de Montpellier (C.G., S.S., L.L., M.K.), Neuroradiology, CHU de Montpellier (N.L.), Department of Pediatric Neurology & Reference Center for Neuromuscular Diseases AOC, CHU Montpellier, PhyMedExp, INSERM, CNRS (F.R., P.M.) – all in Montpellier, France; Université Paris-Saclay, CEA, CNRS, Institute for Integrative Biology of the Cell (I2BC), 91198, Gif-sur-Yvette Cedex, France (F.P-P. and A.D); Translational Stem Cell Biology & Metabolism Program, Research Programs Unit, and Department of Anatomy, Faculty of Medicine, University of Helsinki, 00290, Finland (F.S and T.G.M.); Institute of Child Health, Department of Molecular Neuroscience, University College London (UCL) Institute of Neurology, London, United Kingdom (D.Z and M.R); Radiation Oncology, Albert Einstein College of Medicine, New York, USA (N.M-L.); Institute of Medical Genetics, University of Zurich, Schlieren-Zurich CH-8952, Switzerland (A.B., S.A-B., A.R.); Hertie-Institute for Clinical Brain Research (HIH) and Center of Neurology, and German Center for Neurodegenerative Diseases (DZNE), University of Tübingen, Germany (S.R., L.S., M.S.); Department of Genetics (H.S.A, F.S.A) and Department of Neuroscience (S.A.) King Faisal Specialist Hospital and Research Center,

Riyadh, Saudi Arabia; Dr. John T. Macdonald Foundation, Department of Human Genetics, and John P. Hussman Institute for Human Genomics, University of Miami, Miller School of Medicine, FL33136 Miami, USA (S.Z.); Division of Genetics and Molecular Medicine, King's College London School of Medicine, Guy's Hospital, London, United Kingdom (I.A.B.); Clinical Genetics Unit, Guys and St. Thomas' NHS Foundation Trust, London, United Kingdom (C.D.)

Correspondence should be addressed to Dr. Taylor at the Wellcome Centre for Mitochondrial Research, Translational and Clinical Research Institute, Newcastle University, Framlington Place, Newcastle upon Tyne, NE2 4HH, United Kingdom, or at robert.taylor@ncl.ac.uk

Running title: *ATG7*-linked human disease

Table of contents

3. Supplementary tables

3.1. Table S1: *Clinical features of patients with recessive ATG7 variants*

4. Acknowledgement of infrastructure support

5. References

1. SUPPLEMENTARY METHODS

1.1. Subjects and exome sequencing

Whole exome sequencing of subjects 1 and 2 (family 1) was performed as previously described. Exome capture was acquired using the Illumina TruSeq 62 Mb capture kit, sequenced using the Illumina HiSeq 2000 system in 100 bp reads before alignment to the human reference genome (UCSC hg19). Only variants with a minor allele frequency of ≤ 0.01% from external variant databases gnomAD (http://gnomad.broadinstitute.org/), National Heart, Lung, and Blood Institute Exome Sequencing Project (http://evs.gs.washington.edu/EVS/) and 1000 Genomes Project (http://www.internationalgenome.org/) were considered.

Initially, targeted gene panel sequencing of subjects 3 and 4 was performed by exon capture with the Trusight One sequencing panel kit (www.illumina.com/trusightone) on an Illumina NextSeq sequencer (Montpellier CHRU NGS platform). Subjects 3 and 4, as well as one parent were subsequently WES. Exome capture was acquired using the Twist Human Core Exome Kit (Twist Bioscience, California) and sequenced on an Illumina Novaseq 6000. Reads from the Illumina NovaSeq were processed according to GATK Best Practices recommendations[1,2] and GATK v4.0.3.0 was used unless otherwise stated[3]. Firstly, reads were aligned to the hg38 human reference genome with Burrows-Wheeler Aligner v0.7.15[4] and sorted using SAMtools v1.3.1. Duplicate reads in all lanes were tagged with Picard MarkDuplicates v2.18.2[5]. This was followed by base quality score recalibration, indel realignment, duplicate removal, and SNP and INDEL discovery. Genotyping was performed across all samples simultaneously using standard hard filtering parameters and followed by variant quality score recalibration. Variants passing quality control were annotated with information from external databases using Ensembl Variant Effect Predictor v92.4[6]. Only

variants with a minor allele frequency of ≤ 0.01% from external variant databases gnomAD (http://gnomad.broadinstitute.org/) and 1000 Genomes Project (http://www.internationalgenome.org/) were considered.

Genomic DNA was extracted from EDTA blood of subject 5 and his parents. Single Whole Exome Sequencing (WES) on the index was performed using the xGen® Exome Research Panel v1.0 (IDT) with paired-end sequencing (HiSeq SBS Kit v4, 125 Fwd-125 Rev, Q30-value: 90.1) on a HiSeq2500 sequencer (Illumina Inc.). Raw fastQ files were aligned to the hg19 reference genome using NextGene (Softgenetics). The average depth of coverage was 261x and about 99% of the targeted bases were assessed by ≥20 independent sequence reads (without removing of PCR duplicates). Variants observed in at least 16% of reads with sufficient quality level and with minor allele frequency ≤ 2% (gnomAD) were filtered for in silico gene panels associated with the patients phenotype and further investigated for allele frequency (gnomAD), deleterious in-silico effects assessed by SIFT, Polyphen, LRT, Mutation Taster, Mutation Assessor, FATHMM, M-CAP, GERP and CADD, associations of the affected gene with the patient's phenotype and by literature search for evident functional information. The candidate variants in ATG7 (NM_006395.2) were confirmed in the patient and his parents by Sanger sequencing after PCR amplification using an ABI Genetic Analyzer 3730 (Applied Biosystems, Foster City, California).

In order to validate the association between recessive mutations in *ATG7* and complex neurodevelopmental disorders we filtered the >1500 WES ataxia datasets compiled by the PREPARE ataxia consortium (www.prepare-ataxia.com) within the NGS pipeline GENESIS for recessive, ultra-rare (MAF <0.01%) variants in *ATG7*[7]. This revealed two ATG7 variants in each of the two siblings subject 6 and 7. They had been exome sequenced using the Sure

Select Human All Exon 50MB kit (Agilent) for in-solution enrichment and an HiSeq 2000 (Illumina) to produce about 120 bp paired-end sequence reads. The BWA and Freebayes were used to align sequence reads and call variants. Final data were uploaded into GENESIS pipeline for analysis[7].

For family 5, WES was performed using an Agilent Sureselect All Exons V5 (50 Mb) capture kit (Agilent Technologies; Santa Clara, CA, USA) for library preparation. Briefly, DNA was sheared mechanically after which targeted fragments were captured by probe hybridization and amplified before sequencing. An Illumina HiSeq 2500 (Illumina Inc; San Diego, CA, USA) was used for paired-end 100 nt sequencing. Sequence alignment, indexing of the reference genome (hg19), variant calling and annotation used a pipeline based on BWA, Samtools, GATK and Annovar, respectively. Essentially, variants were annotated using a combination of public knowledge databases available from the Annovar package and in-house databases which included collections of previously published Saudi disease causing variants. Autozygome definition solely from whole exome sequence data was undertaken as previously described[8].

1.2. Reverse transcription and cDNA Sequencing

Whole blood from patients and controls was collected into PAXgene™ Blood RNA System tubes (PreAnalytiX, Qiagen) and total mRNA was isolated using PAXgene™ reagents according to the manufacturer's protocol. Concentration and purity were determined by the NanoDrop ND1000 spectrophotometer (Peqlab) and 500 ng RNA were used for reverse transcription with the Transcriptor High Fidelity cDNA Synthesis Kit (Roche). To identify splice variants specific primers amplifying *ATG7* coding exon 5 to exon 10 (F: TGCCCTCTGTCTTCCAGAGA; R: TACATTGCAACCCAAGGTGC) were used with an

annealing temperature of 56°C, elongation time of 90 seconds. Distinct PCR bands were sequenced with the BigDye Terminator v3.1 Cycle Sequencing Kit and the automated sequencer ABI Prism 3130xl Genetic Analyzer (Applied Biosystems). Sequences were visualized with Staden 2.0.0b10 software (Staden Sourceforge).

1.3. RNA sequencing and analysis

RNA-sequencing was performed on a total of 58 individuals. This comprised of 5 *ATG7* patients and an additional 53 patient samples with varied, unrelated phenotypes that were used as controls. RNA from fibroblast pellets was isolated using the miRNeasy 96 sample kit (Qiagen, UK) in accordance with the manufacturer's protocol. The RNA integrity number was assessed for each sample using the Agilent 2100 Bioanalyzer (Agilent Technologies UK Ltd, UK) in combination with the RNA 6000 Nano-LabChip kit, with all RIN values found to exceed 8.0. Using 100 ng of total RNA as input, cDNA libraries was prepared using the Illumina TruSeq Stranded mRNA System v2 (Illumina, US) according to the manufacturer's protocol. Paired-end sequencing was performed on the Illumina NovaSeq 6000 Sequencing System to generate ~100 million pair-end reads per sample. Reads of 150-bp length were used to increase the proportion of reads with a gapped alignment, which can represent splicing events. Pre-alignment quality control including adapter trimming, read filtering and base correction was performed using fastp (v0.20.0)[9]. Reads were aligned using STAR 2-pass (v2.7.0) to the hg38 build of the reference human genome (hg38) using gene annotation from Ensembl v97[10]. 1st pass alignment was used to discover novel junctions, which were used as input to the 2nd pass to improve the sensitivity of junction detection. Reads were required to be uniquely mapped to only a single position in the genome. The minimum required overhang length of an annotated and unannotated junction was required to be 8 and 3 base

pairs respectively. The output BAM files underwent post-alignment QC using RSeQC (v2.6.4) with all samples passing quality control after manual assessment[11].

dasper (v1.0.0) was used to detect aberrant splicing events in each of the 5 ATG7 patients. BAM files were first converted into BigWig format using megadepth. Junctions outputted by STAR and BigWig files were used as input into the dasper pipeline. dasper was used to filter junction data, for which we required junctions to have at least 5 counts in at least 1 sample, with an implied intron length between 20-1,000,000 base pairs and to be outside ENCODE blacklist regions[12]. Junctions were then annotated using Ensembl v97 reference annotation and clustered by their shared acceptor or donor sites with their counts normalised to permit comparison between patients and controls. Coverage data across the flanking exons and intron associated with each junction was calculated and normalised for each sample. Junction and coverage normalised counts were scored using the default z-score approach provided by dasper. Resulting z-scores for each patient were placed into an isolation forest model to rank splicing events by their aberrance in relation to controls. Sashimi plots describing the splicing in Subject 1 was plotted using dasper.

Gene count matrices were generated using RNA-SeQC (v2.3.4) using Ensembl v97 reference annotation[13]. Only genes with an FPKM of 1 in at least 5% of samples were retained for downstream analyses. OUTRIDER (v1.6.1) was applied with default settings to find genes with outlier expression in each ATG7 patient. A 1-vs-all experimental design was used whereby each ATG7 patient was compared to the remaining 53 controls[14]. The OUTRIDER autoencoder was fitted using the optimal encoding dimension found to the be 12. The Benjamini-Hochberg method for multiple testing correction was applied to p-values, with a significance threshold of 0.05.

DESeq2 (v1.26.0) was used to detect common genes or pathways that were dysregulated across all *ATG7* patients[15]. Thus, all 5 *ATG7* samples were grouped together as the case cohort and compared to the 53 control samples. The Benjamini-Hochberg method was used for multiple test correction and genes with a p-value less than 0.05 and an absolute log-fold change greater than 1.5 were considered differentially expressed. gprofiler2 (v0.1.9) was then run with differentially expressed genes ranked by p-value as input, to find significantly dysregulated pathways[16]. As the background gene list for gprofiler2, we used all genes that passed the expression filter.

1.4. Transmission Electron Microscopy

Samples were collected and fixed in 2% glutaraldehyde in sodium cacodylate buffer at 4 ^{o}C then 1% osmium tetroxide. Samples were dehydrated using graded acetone then impregnated with graded resin over 24 hours before embedding in 100% resin at 60 ^{o}C for 24 hours. Ultrathin sections approximately 70 nm were cut, stretched with chloroform and mounted on Pioloform-filmed copper grids. Sections were stained using 2% aqueous uranyl acetate and lead citrate. Grids were imaged on a Hitachi HT7800 transmission electron microscope using an Emsis Xarosa camera with Radius software. For quantification of autophagy in primary fibroblasts, 4000X images were taken and analysed under blinded conditions. The area of autophagic structures were estimated for each cell using FIJI. Cytoplasmic area was calculated by removing nuclear area from total cell area.

1.5. Cell culture

Primary and immortalised (*via* hTERT activation) dermal fibroblasts, and mouse embryonic fibroblasts, were maintained in DMEM (Gibco, 41966052) supplemented with 10% FCS, 50

μg.ml^{-1} uridine, 1X non-essential amino acids and 1X penicillin/streptomycin and incubated at 37 °C in 5% humidified CO_2. Myoblasts were grown in Skeletal Muscle Cell Growth Medium with supplement mix (PromoCell, C-23060), and the addition of 10% FCS, 1X GlutaMAX (Gibco, 35050-038) and 80 ug.ml^{-1} gentamycin (Gibco, 15710-064).

1.6. Immunoblotting

Cells were lysed with lysis buffer (50mM Tris/HCl pH 7.4, 130 mM NaCl, 2 mM $MgCl_2$, 1 mM phenylmethylsulfonyl fluoride (PMSF), 1% Nonidet P-40, 1X Complete EDTA-free protease inhibitor (Roche)) prior to resolution *via* centrifugation at 5,000 *g*. Lysates were analysed *via* SDS-PAGE and immunoblotting. 10-30 μg of protein was loaded, and SDS-polyacrylamide gels contained 12% bis-acrylamide unless stated otherwise.

The following primary antibodies were used, diluted 1:1000 unless stated otherwise: ATG7 (Millipore 04-1055/Abcam ab52472), ATG3 (1:500, Cell Signaling #3415) ATG5 (1:500, Cell Signaling #2630), P62 (Abcam ab109012), ULK1 (Cell Signaling #8054), LC3B (Cell Signaling #2775), β-actin (1:10000, Cloud Clone Corp. CAB340Hu22) and GAPDH (1:10000, Proteintech 60004-1-Ig). Anti-mouse (1:2000, DAKO P0260) and anti-rabbit (1:3000, DAKO P0399) secondary antibodies were used. Images were analysed and processed using Image Lab (BioRad).

1.7. Immunofluorescence

Immortalised fibroblasts were seeded on μ-slide 8-well chambered coverslips (Ibidi). Cells were fixed in 100% methanol for 5 minutes, permeabilised and blocked (1% BSA, 0.1% Triton X-100, 10% normal goat serum) and then stained using the following primary and secondary antibodies: P62 (1 μg/ml, Abcam ab109012), Ubiquitin (1:500, Abcam ab7254), goat anti-mouse IgG1 AlexaFluor 488 (1:200, Invitrogen A-21121) and goat anti-rabbit IgG

AlexaFluor 647 (1:200, Invitrogen A32733). Cells were imaged using an oil-immersion 100X objective on VisiTech iSIM array-scanning confocal microscopy associated with an inverted Nikon TiE body. A 488 nm laser (100%) and 642 nm laser (3%) were used with a 700 ms exposure. Z-stacks were deconvolved using Huygens Software (automated; +5% for P62 channel, and +25% for ubiquitin channel) before thresholding and analysis via Objects Analyser on Huygens Software (Scientific Volume Imaging). Threshold and seed were both set at 3,000, with garbage volume set at 5 voxels. P62 puncta were defined by their number and volume. Cell size was measured manually using FIJI.

1.8. Autophagic flux assays

Fibroblasts were seeded overnight to approximately 60% confluence then, after replacement of media, treated with 60 μM chloroquine (Sigma) or 100 nM Bafilomycin A1 (Enzo Life Sciences) for 2 hours to block autophagic flux, unless otherwise stated. Autophagy was induced using 1 μM AZD8055 (Santa Cruz) or acute serum and amino acid removal using Earle's Balance Salt Solution (EBSS).

1.9. Autophagic sequestration assay

The long-lived protein degradation (LLPD) assay (autophagic sequestration assay) was performed by activity measurement of autophagosomal lactate dehydrogenase (LDH) as described previously[17,18]. Patient-derived cells were cultured at 37°C in DMEM with high glucose supplemented with 1X GlutaMax, 10% FBS and 100U/ml Penicillin-Streptomycin (Life Technologies/Gibco). Non-selective macroautophagy was induced in cultures via acute serum and amino acid starvation by EBSS and addition of 100 nM Bafilomycin A1 (BafA1) in EBSS medium for 3 hours at 37°C at 5% CO_2. Cells were harvested in 0.25% trypsin-EDTA and collected in PBS containing 5% BSA. Following centrifugation at 500 x *g* for 5

minutes at 4°C, we resuspended cells in ice-cold 10% sucrose. Plasma membrane disruption was performed following published recommendations, *via* electroporation (BioRad) at 800 V, 25 μF, and 400 Ω. After electrodisruption, we resuspended cells in 400 μl of phosphate-buffered sucrose (100 mM sodium monophosphate, 2 mM EDTA, 2 mM DTT and 1.75% sucrose, pH 7.5). 550 μl from the disrupted cell suspension solution was resuspended in 900 μl ice-cold resuspension buffer (50 mM sodium monophosphate, 1 mM EDTA, 1 mM DTT) for the of the sedimented measurement (LDH $_{\text{"Sediment"}}$). Autophagic vacuoles were sedimented by centrifugation at 18,000 x *g* for 45 minutes at 4°C. The supernatant was aspirated, flash-frozen with liquid nitrogen and sored at -80°C. We collected 150 μl from the disrupted cell suspension solution for the total LDH measurement, flash-frozen with liquid nitrogen and stored at -80°C (LDH $_{\text{"Total"}}$). LDH $_{\text{"Sediment"}}$ was diluted in resuspension buffer with 1% and LDH $_{\text{"Total"}}$ with 1.5% Triton X-405 and rotated in a cold room at 4°C for 30 min. After centrifugation at 18,000 x *g* for 5 minutes at 4°C, we measured LDH activity as described previously in a working solution containing 0.6 mM pyruvate and 0.36 M NADH[1,2].

1.10. DNA transfections

pcDNA3.1(+) plasmids encoding wildtype *ATG7* (CDS derived from NP_006386.1), catalytic null ATG7 (C572A) or missense variants (R576H, H624Y, P234T, V588M, G511D) were generated by GenScript. Immortalised fibroblasts (approximately 120,000) were seeded to 70-80% confluence overnight in 12-well plates, then incubated with 500 ng of plasmid encoding wildtype ATG7 and 1.5 µl FuGene HD transfection reagent, for 24h, before media was replaced and cells were treated with or without 60 µM chloroquine and 1 µM AZD8055 for two hours. Plasmid and transfection reagent were prepared in Opti-MEM (Gibco). For transfection of *Atg7* knockout (KO) MEFs, which were kindly acquired from Dr

Viktor Korulchuk with permission from Professor Masaaki Komatsu, 30,000 *Atg7* KO MEFs were seeded overnight in 12-well plates then incubated for 24h with 500 ng of plasmid encoding ATG7 with 1.5 μl FuGene HD transfection reagent (Promega). Plasmid and transfection reagent were prepared in Opti-MEM (Gibco). After 24h incubation, media was replaced to remove plasmid and transfection reagent, before treating cells with 60 μM chloroquine and 1 μM AZD8055 for two hours. For all transfections, PenStrep was removed from media.

1.11. Yeast studies

Yeast strains were grown at 28˚C in synthetic medium with glucose (SD; 0.67% yeast nitrogen base without amino acids and 2% glucose). Whenever necessary, media were supplemented with the appropriate nutritional requirements according to the strains. SD(-N) (2% glucose, 0.17% yeast nitrogen base without ammonium sulfate) was used as starvation medium for the autophagy assays. To induce autophagy, fresh cultures were diluted to OD600 = 0.03 in SD and incubated at 28°C for 14 hours to mid-log phase before they were shifted to SD(-N) for nitrogen starvation during appropriate time period.

Standard genetic techniques were applied for the growth and manipulation of *Saccharomyces cerevisiae* cells. The strain BY4741 deleted for *atg7* (*atg7Δ*: MATa, *his3-Δ1 leu2-Δ0 met15-Δ0 ura3-Δ0 atg7Δ::*KAN) was obtained from the EUROSCARF consortium. The *pho8Δ60 atg7Δ* isogenic double mutant strain was obtained after crossing *atg7Δ* with the BY4742 *pho8Δ60* strain and spore dissection. Transformation of yeast strains was performed by the Lithium acetate method.

The *ATG7* gene with its promoter and terminator sequences was amplified by PCR, sequenced and cloned into the low copy plasmid pFL38(*URA3*) at the BamHI and PstI sites. The *ATG7* plasmids carrying each point mutation were generated from pFL38-*ATG7* through a site directed mutagenesis using the Phusion™ Site-Directed Mutagenesis Kit protocol (Thermoscientific) according to the manufacturer's recommendations. Direct sequencing was performed to confirm the successful introduction of each mutation. The wild type *ATG7* and mutant sequences were introduced into the pFL36(*LEU2*) plasmid by direct cloning (BamHI and PstI) from the pFL38 constructions. The oligonucleotide sequences used for all these constructions are as follows:

	Oligo sequence 5' → 3'
ATG7 cloning	CGGGATCCGTTAACCTATACAAAAACCGC
ATG7 cloning	AACTGCAGCAATGTTATAAACGTTAGAGTC
R511H mutagenesis	CAAATGTGCACAGTAACTCACCCCGGCGTTGCTATGATG
R511H mutagenesis	ATCCAAAGTCCTGTCAGTTAAACTGTCAGTTGG
H547Y mutagenesis	GTATTAGGGGATATACCTTATCAAATACGTGGGTTTTTG
H547Y mutagenesis	TGTTGTTTCTGAACCAGAGTATTTGGT
T208A mutagenesis	GTTTTGGCGATAAGGGACGCTAGTACGATGGAAAACGTC
T208A mutagenesis	TTTAGTTTTTCGAATAATACACTTGTCATAATTTAC
V523M mutagenesis	GCCTCTTCTTTAGCAATGGAATTGATGACTTCC
V523M mutagenesis	CATCATAGCAACGCCGGGTCTAGTTAC
G473D mutagenesis	CTTAGTTATGAGGCATGATAATAGAGATGAACAG
G473D mutagenesis	TAGCTATCAAACCCCAGAGCAGCATTAAT

GFP-Atg8 assay

The *atg7Δ* strains expressing ATG7 or its mutated forms were transformed with a LEU2 plasmid allowing expression of GFP-Atg8 from its own promoter (kind gift from N.

Camougrand). Logarithmic growing cells in SD were transferred to the starvation medium SD(-N) for the indicated time and total protein extraction was performed using the procedure depicted in Kainz et al., 2016. Total protein extracts were loaded on 12% sodium dodecyl sulphate–polyacrylamide gel electrophoresis gels and transferred to nitrocellulose membranes by wet transfer. Membranes were probed with monoclonal anti-GFP antibody (Invitrogen). Immunodetection was performed with horseradish peroxidase-conjugated anti-mouse antibodies using chemoluminescence western blotting reagents (Pierce) according to the manufacturer's instructions. Phosphoglycerate kinase (Pgk) was detected using the mouse anti-Pgk (Invitrogen) and used as a loading control. Quantification was performed using ImageLab version 6.1.1 Build 7 © 2020, Bio-Rad Laboratories, Inc. software.

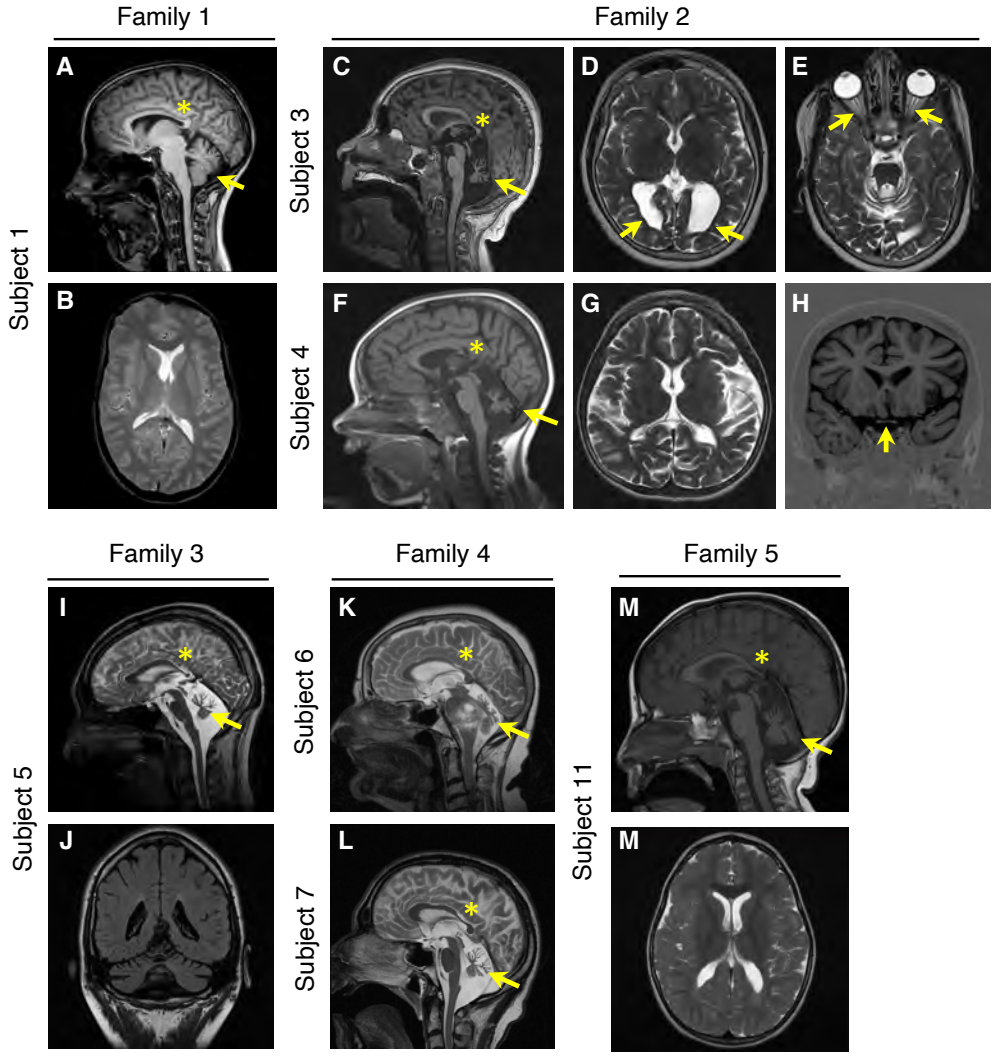
Pho8Δ60 assay

Cells were grown to mid-log phase, switched to nitrogen-starvation media (SD(-N) for 6 h to induce autophagy and the Pho8Δ60 assay was carried out as previously described (Noda and Klionsky, 2008). At each time point, 3 OD_{600} units of cells were collected by centrifugation at 1,500 x g for 10 min. The cell pellet was then washed in H2O and resuspended in 200 µL ice-cold assay buffer (250 mM Tris-HCl, pH 9,0; 10 mM MgSO4; 10 µM ZnSO4) and kept on ice. Acid-washed glass beads (0.2 ml) were then added and the samples were vortexed at 4°C for 4 min (in 1 min intervals). The lysed cells were pelleted at 14,000 x g for 10 min at 4°C and the supernatant was collected. Meanwhile, the α-naphthyl phosphate disodium salt substrate (α-NP; Sigma N7255) was diluted to 55 mM in assay buffer. The assay was carried out by mixing 50µl 55mM α-NP with 50 µl cell lysate in 500 µl assay buffer at 30°C. For each condition, the reaction was done in duplicate and incubated at 30° C for 20 min. The reaction was stopped by adding 0,4 ml of 2M Glycine-NaOH, pH11,0. The fluorescence was measured using a wavelength of 345 nm for excitation and 472 nm for emission (Plate reader Infinite 200

PRO, TECAN). The BCA assay was carried out to determine protein concentration of each lysate sample, and enzyme activity was calculated. Three independent experiments were carried out and results reported are mean ± SD.

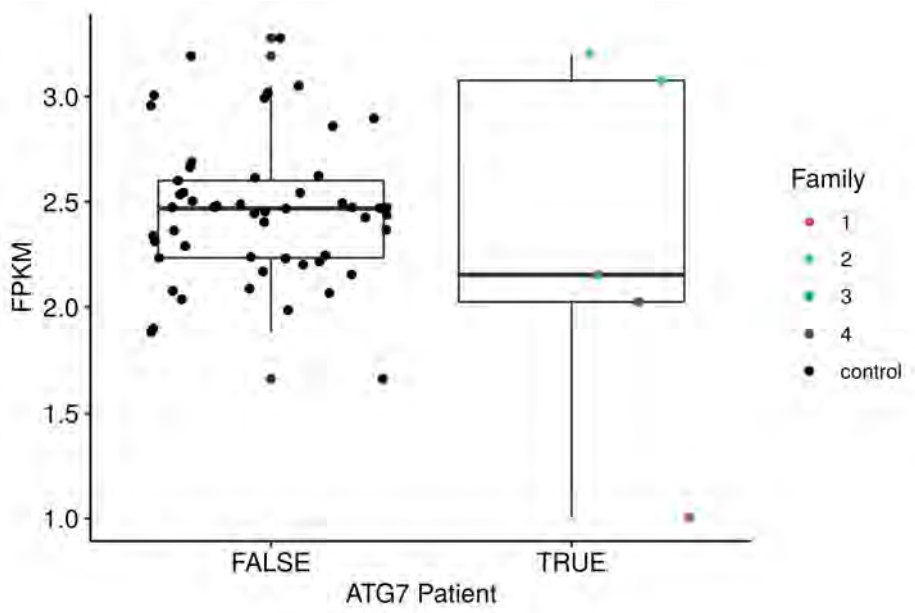
2. SUPPLEMENTARY FIGURES

2.1. Figure S1



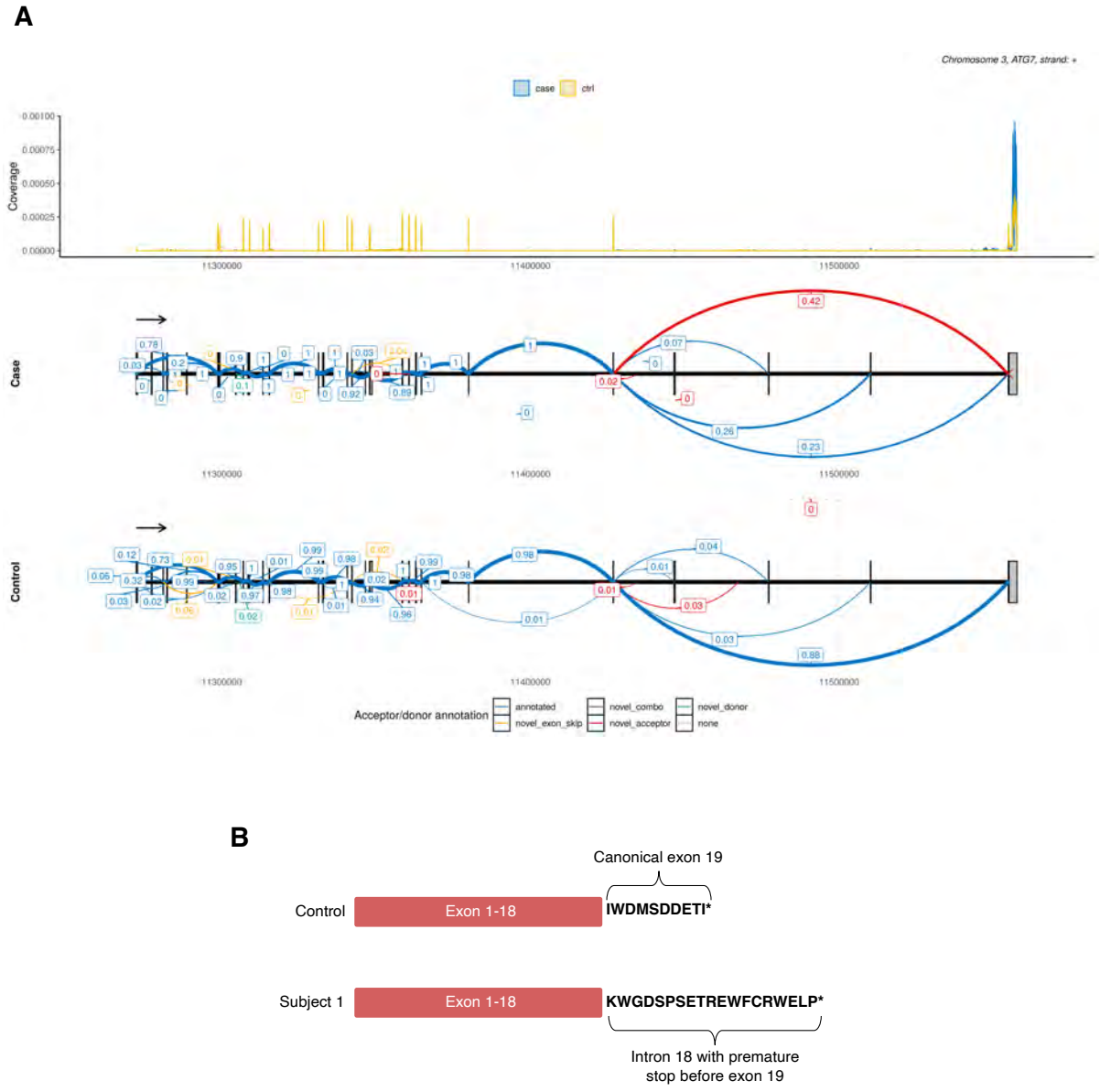
Patient magnetic resonance imaging demonstrates brain abnormalities. On median sagittal T1-weighted sections (**A, C, F, I, K, L, M**): pontocerebellar hypoplasia involving predominantly the vermis (arrow); posterior atrophy of the corpus callosum (*). On axial T2-weighted slides (**D**): enlarged posterior lateral ventricles secondary to loss of white matter (arrows) with further evidence of this process involving the cerebral hemispheres on image (**G**). On axial T2-weighted (**E**) and on coronal T1-weighted sections (**H**): bilateral atrophy of the optic nerves and Chiasma (arrow).

2.2. Figure S2



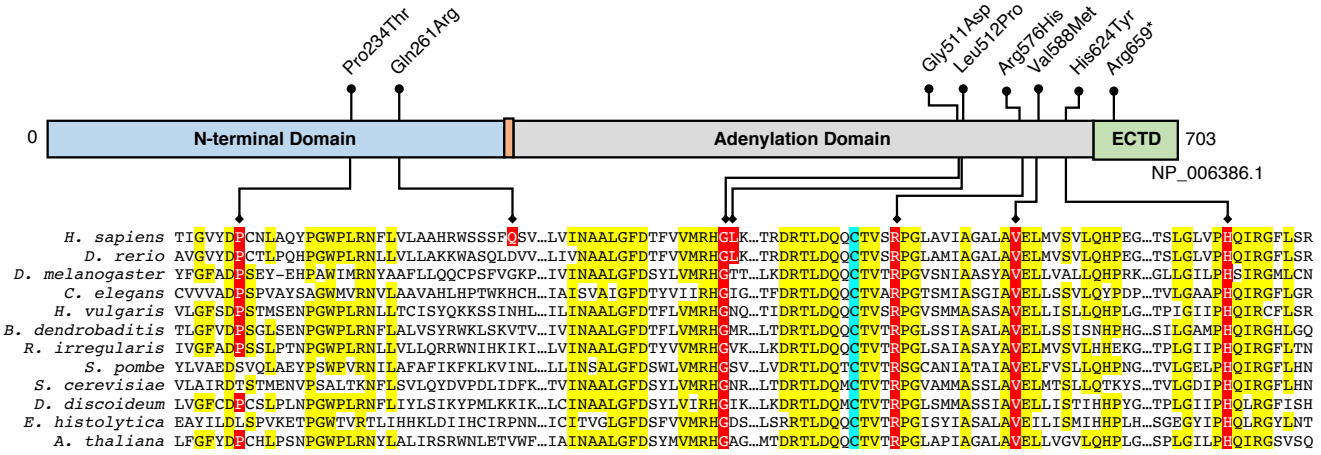
Analysis of ATG7 expression via RNA sequencing. *FPKM = Fragments per kilobase million.*

2.3. Figure S3



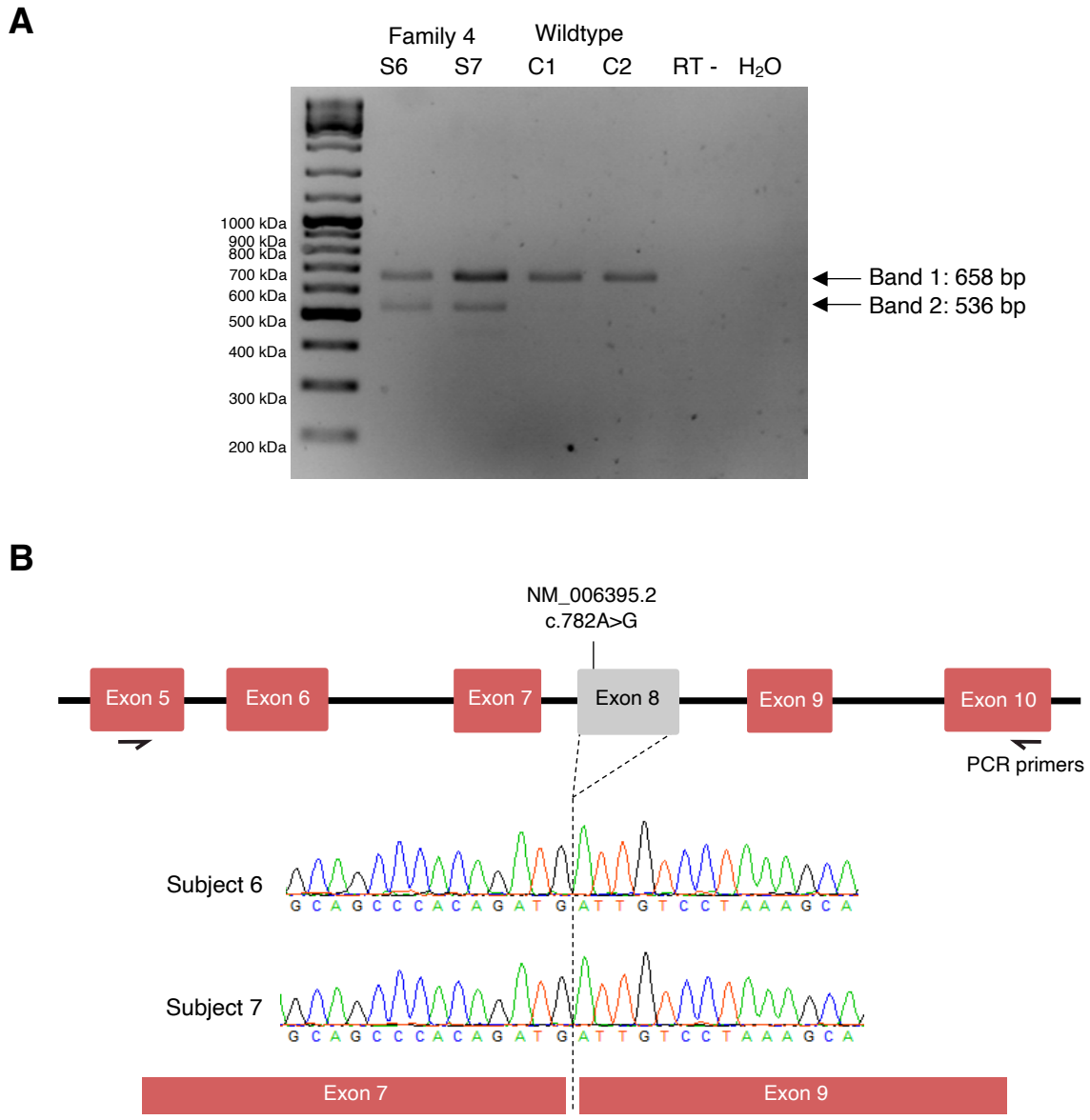
Analysis of c.2080-2A>G variant. **(A)** Sashimi plot describing the splicing across *ATG7* in control and case (subject 1) using *dasper*. **(B)** Schematic diagram illustrating the replacement of canonical exon 19 with intron 18, including a premature termination, resulting from the c.2080-2A>G *ATG7* variant affecting family 1.

2.4. Figure S4



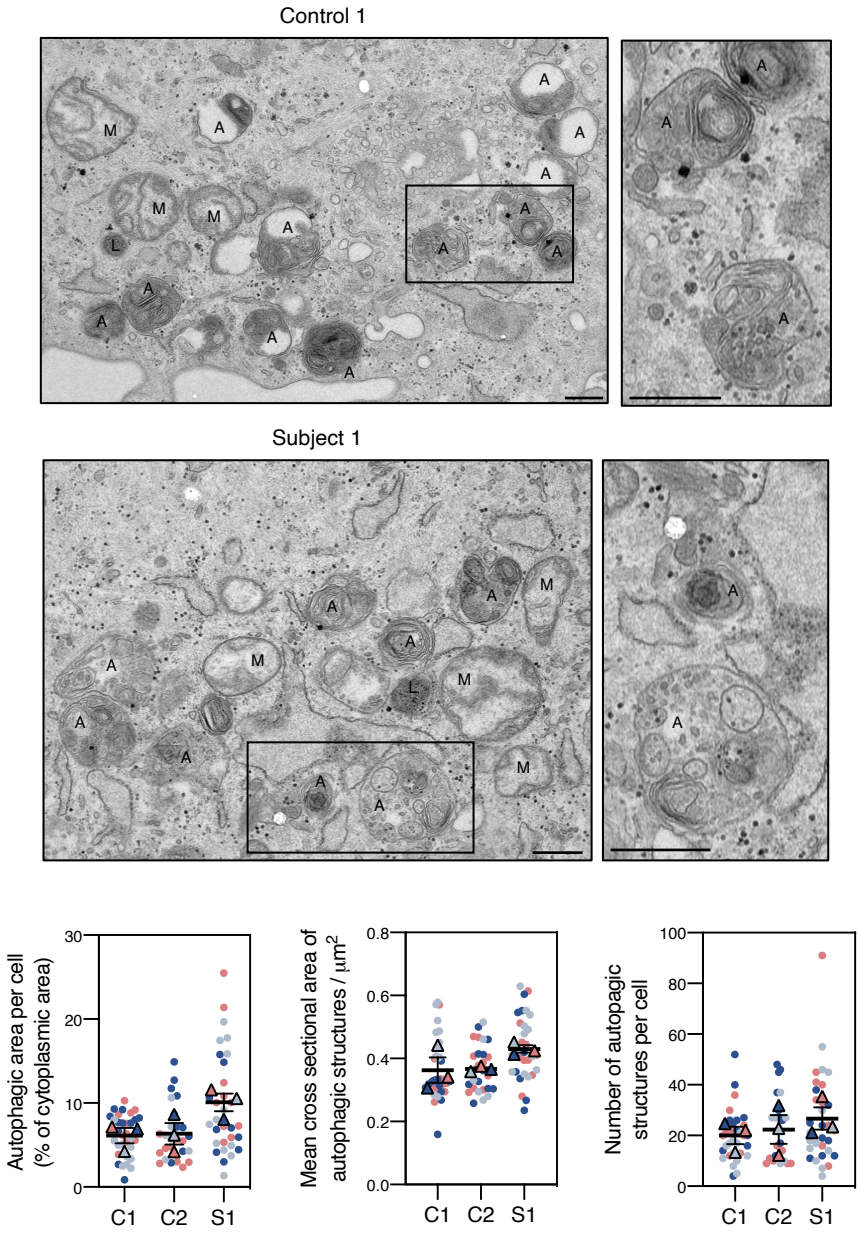
Conservation of amino acids. Amino acid sequence comparison of human ATG7 (NP_006386.1) and orthologous proteins from different species highlights the conservation of affected amino acids in *ATG7* patients. Highly conserved amino acids are highlighted in yellow, red (mutated in families 2-5) or blue (catalytic cysteine 572). ECTD = extreme c-terminal domain.

2.4. Figure S5



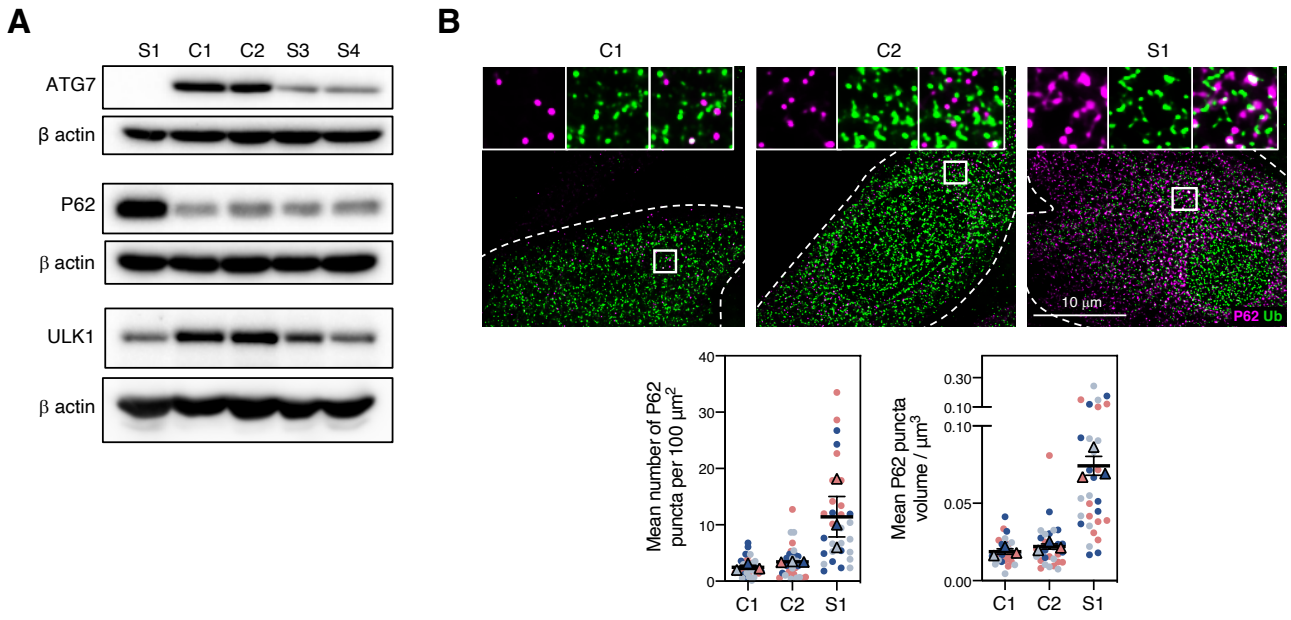
Analysis of c.782A>G variant. (**A**) PCR of *ATG7* cDNA spanning coding exon 5 to exon 10 showed an additional PCR product for patients S6 and S7 of family 4 compared to two controls. (**B**) Sequencing of the distinct PCR bands revealed the loss of coding exon 8 (122 bp) in patients harboring the NM_006395.2: c.782A>G mutation. bp, basepair; kDa, kiloDalton; RT(-), cDNA transcription without reverse transcriptase.

2.4. Figure S6



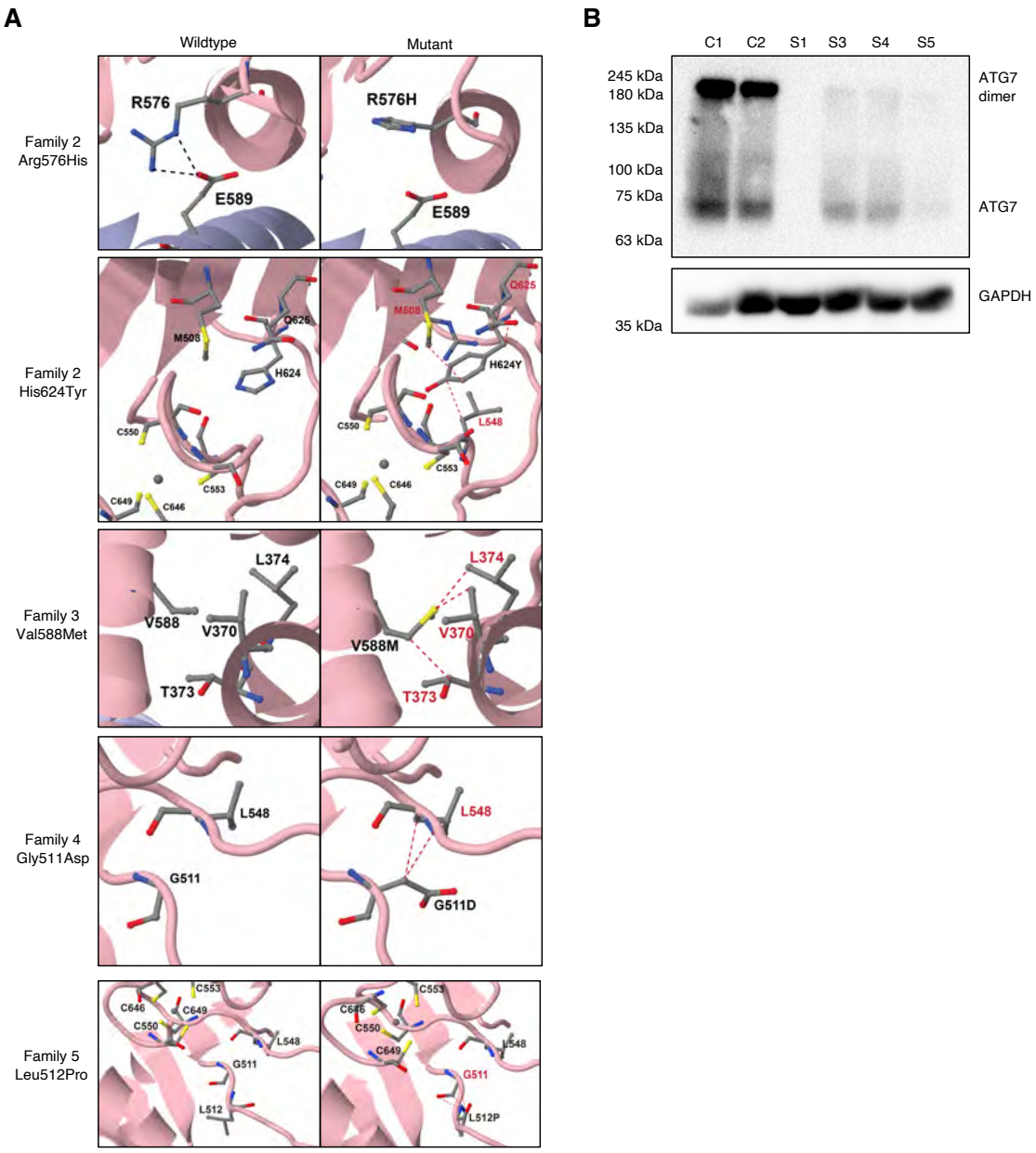
Transmission electron microscopy (TEM) analysis of Subject 1 fibroblasts. Blinded analysis of TEM imaging of primary fibroblasts treated with 100 nM bafilomycin A1 for 2 hours. At least 26 cells in total per cell line, across 3 independent repeats, were randomly imaged. Data shows mean values +/- SEM. Representative images highlight autophagic structures from control 1 and Subject 1 (scale bars = 500 nm). A = autophagic structures; L = lysosome; M = mitochondrion. Circles represent mean values for an individual image field; triangles represent mean values for each independent biological repeat; blue, grey and pink represent biological repeats.

2.7. Figure S7



Analysis of immortalised fibroblasts via western blotting and immunofluorescence. **(A)** Western blot analyses of autophagy-related proteins using antibodies reactive to ATG7, P62, and ULK1 in immortalised patient fibroblasts. ATG7 was undetectable in Subject 1 and diminished in Subjects 3 and 4. P62 steady-state levels were elevated in subject 1, and ULK1 deficiency was identified in all patient cell lines, most notably in Subject 1. β actin was used as a loading control. **(B)** Immortalised fibroblasts, cultured under basal conditions, were methanol-fixed and imaged *via* confocal microscopy after staining using antibodies against P62 and ubiquitin (Ub). 30 images acquired across 3 independent experiments were analysed using Huygens Imaging Software. Representative images are illustrated. Graphs illustrate the number and volume of P62 puncta in samples analysed. Data represents mean values +/- SEM. Circles represent mean values for an individual image field; triangles represent mean values for each independent biological repeat; blue, grey and pink represent biological repeats.

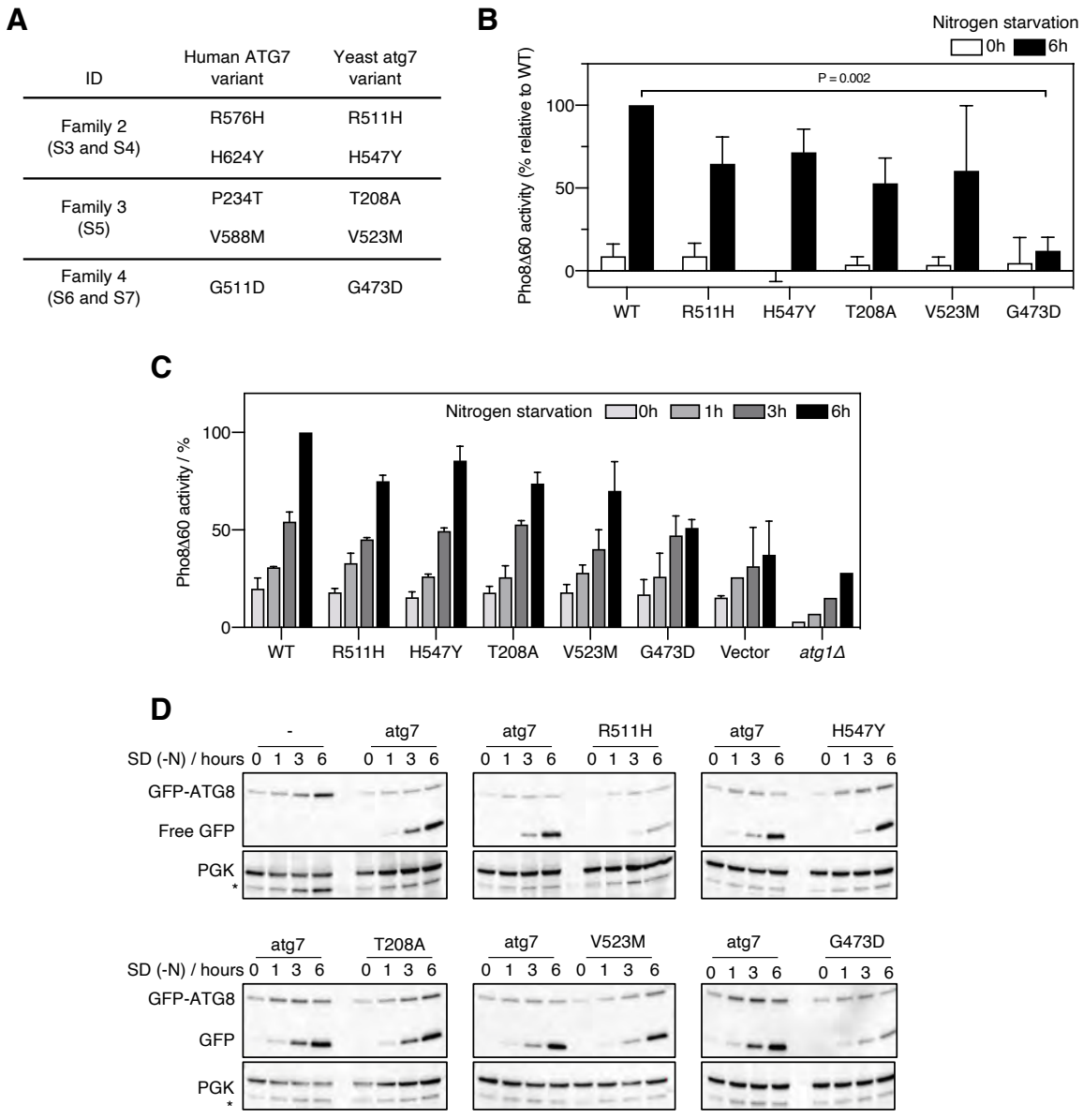
2.4. Figure S8



In silico modelling of missense ATG7 variants. **(A)** Homology models of the human ATG7 homodimer (monomer 1 coloured in pink, monomer 2 in purple) built from the crystal structure of the yeast Atg7-Atg3 complex (PDB: 4GSL (Arg576, Val588 and His624) and the Atg7-Atg8 complex (PDB: 3hv3 (Gly511)). Arg576, Val588 and His624 are located in the conserved homodimeric adenylation domain of ATG7. Arg576His induces a loss of a salt bridge (black dotted lines) between Arg576 and Glu589 in wildtype; Val588Met induces steric clashes with residues Val370, Thr373 and Leu374 (labelled in red) which might

prevent optimal folding and interfere with dimerization. His624Tyr is predicted to induce steric clashes with residues M508, L548 and Q625 (labelled in red), disturbing the region which includes a zinc finger, preventing optimal folding of the domain. Gly511Asp is predicted to induce a steric clash with the residue L548 (labeled in red). Putative clashes are represented by red dotted lines. Pro234 is located in the N-terminal domain, which does not display sequence homology between yeast and human, so models could not be generated. **(B)** Non-reducing SDS-PAGE (10% bis-acrylamide) was undertaken before immunoblotting with an antibody reactive to ATG7 in primary fibroblasts derived from subjects 1, 3, 4 and 5 and age-matched controls.

2.5. Figure S9



***Yeast complementation experiments*. (A)** Table showing the homologous yeast *ATG7* variants expressed in *atg7* KO strains. **(B)** Autophagy was induced in *atg7* KO yeast expressing *pho8*Δ60 and homologous *atg7* variants *via* nitrogen starvation for up to 6 hours before spectrophotometric analysis of Pho8Δ60 activity – a measure of autophagic activity. *Atg7* KO yeast expressing a control vector were also studied. For analysis shown, Pho8Δ60 activity from control vector-expressing yeast (i.e. background) was subtracted from activity

measured in yeast expressing wild type and mutant *atg7*. All values were made relative to WT values at 6 hours. Data show mean values + SD. Adjusted P values are based on a Kruskal-Wallis test and Dunn's multiple comparisons to WT values. Data were derived from four independent experiments. **(C)** Pho8Δ60 activity was measured in an *atg7* knockout strain of *S. cerevisiae* expressing variant *atg7* and starved by removal of nitrogen for up to 6h. Data were normalized to WT values at 6h. Mean values are shown + SD. **(D)** Homologous *atg7* mutants were expressed in an *atg7* knockout strain of *S. cerevisiae* expressing GFP-atg8 before nitrogen starvation for up to 6 hours. Representative western blots (n = 3) are shown for each variant assessed after detection of GFP and PGK (loading control).

3. SUPPLEMENTARY TABLE

3.1. Table S1

	Family 1		Family 2		Family 3	Family 4	
	Subject 1	**Subject 2**	**Subject 3**	**Subject 4**	**Subject 5**	**Subject 6**	**Subject 7**
Age	28 years	18 years	18 years	15 years	34 years	71 years	68 years
Sex	Female	Female	Female	Female	Male	Female	Male
Age at last investigation	28 years	18 years	18 years	15 years	34 years	71 years	58 years
Geographic origin	United Kingdom	United Kingdom	Morocco & Germany	Morocco & Germany	Switzerland	Germany	Germany
Mutations* (protein alteration)	c.1975C>T (p. Arg659*) c.2080-2A>G	c.1975C>T (p. Arg659*) c.2080-2A>G	c.1727G>A (p.Arg576His) c.1870C>T (p.His624Tyr)	c.1727G>A (p.Arg576His) c.1870C>T (p.His624Tyr)	c.700C>A (p.Pro234Thr) c.1762G>A (p.Val588Met)	c.782A>G (p.Gln261Arg) c.1532G>A(p.Gly511Asp)	c.782A>G (p.Gln261Arg) c.1532G>A(p.Gly511Asp)
Method of mutation detection	WES	WES	WES	WES	WES	WES	WES
Birth weight	3.1 kg	3.1 kg	NC	NC	NC	NC	NC
Age of walking	22 months	18 months	Never achieved	Never achieved	18-24 months	NC	NC
Latest Head circumference centile / cm	10th	2nd	51 cm	48 cm	25-50th	NC	NC
Weight / height	40.3 kg / 157 cm	37.2 kg / 159 cm	37 kg / 153 cm	19 kg / 121 cm	76 kg / 187 cm	NC	NC
Growth retardation	NC	NC	+	+	-	NC	NC
Seizures (age of onset)	-	-	Tonic clonic seizures with cyanosis (7y) and status epilepticus (15y)	-	-	-	-
Intellectual disability	Moderate	Mild	Severe	Severe	Moderate	Mild	Moderate
Early-onset encephalopathy	-	-	+	+	-	-	-
Neurological examination							
Abnormal Tone	-	-	Truncal hypotonia with spastic paraplegia	Truncal hypotonia with spastic paraplegia	-	-	-
Muscle weakness	+	+	-	-	+	-	-
Movement disorder	Ataxia; tremor	Ataxia; tremor	Ataxia; tremor; dyskinesia	Ataxia	Ataxia	Ataxia; tremor	Ataxia; dyskinesia
Eye abnormalities	OA; CPEO; ptosis	OA; CPEO; ptosis; early-onset cataracts;	Retinopathy	Retinopathy; wandering; amblyopia; OA	-	ND	-
Speech	Dysarthria	Nasal speech	No language	No language	Normal	Normal	Basic vocabulary
Deafness	+ (Bilateral SNHL)	+ (Bilateral SNHL)	-	-	-	-	-
Further clinical observations							
Clinodactyly of 5th fingers	+	+	-	-	+	-	-
Other skeletal abnormalities	Pectus excavatum	Pectus excavatum	NC	Lumbar kyphoscoliosis	NC	-	-
Facial dysmorphism	High arched palate; gum hypertrophy; long narrow face; retrognathia	High arched palate; gum hypertrophy; long narrow face; retrognathia	-	-	High arched palate Smooth philtrum	-	-
Other	-	Autistic spectrum disorder	EEG abnormalities	Phospho-calcium metabolic disorder; severe feeding difficulties; EEG abnormalities	Pes cavus; diarrhea and no appetite; hypogonadotropic hypogonadism; hypertrophic cardiomyopathy; gynecomastia	Late onset-dementia decline; schizophrenic psychosis	Generalised choreatic hyperkinesia; bruxism; dystonia; aggression; self-mutilation behaviour; severe dysphagia
Neuroimaging							
Brain MRI abnormalities	Cerebellar hypoplasia; Posterior atrophy of corpus callosum	NA	Cerebellar hypoplasia; posterior atrophy of corpus callosum; bilateral opto-chiasmatic atrophy; ventriculomegaly predominant in the occipital horns due to rarefaction of white matter	Cerebellar hypoplasia; posterior atrophy of corpus callosum; bilateral opto-chiasmatic atrophy; atrophy of the two cerebral hemispheres with rarefaction of white matter	Cerebellar hypoplasia; posterior atrophy of corpus callosum	Cerebellar hypoplasia; posterior atrophy of corpus callosum	Cerebellar hypoplasia; posterior atrophy of corpus callosum

**Nomenclature HGVS V2.0 according to mRNA reference sequence NM_006395.2. Nucleotide numbering uses +1 as the A of the ATG translation initiation codon in the reference sequence, with the initiation codon as codon 1.* OA, optic atrophy; CPEO, chronic progressive external ophthalmoplegia; MRI, magnetic resonance imaging; SNHL, sensorineural hearing loss; NA, not analysed; NC, not communicated; ND, not determined

	Family 5				
	Subject 8	**Subject 9**	**Subject 10**	**Subject 11**	**Subject 12**
Age	21 months	16 years	14 years	5 years	4 years
Sex	Female	Female	Male	Female	Female
Age at last investigation	21 months	16 years	14 years	5 years	4 years
Geographic origin	Saudi Arabia	Saudi Arabia	Saudi Arabia	Saudi Arabia	Saudi Arabia
Mutations* (protein alteration)	Patient was not sequenced	c.1535T>C (p.Leu512Pro) c.1535T>C (p.Leu512Pro)	c.1535T>C (p.Leu512Pro) c.1535T>C (p.Leu512Pro)	c.1535T>C (p.Leu512Pro) c.1535T>C (p.Leu512Pro)	c.1535T>C (p.Leu512Pro) c.1535T>C (p.Leu512Pro)
Method of mutation detection	NA	WES	WES	WES	WES
Birth weight	ND	ND	ND	3 kg	3 kg
Age of walking	Never achieved	Never achieved	9 years	3 years	3.5 years
Latest Head circumference centile / cm	NC	NC	NC	48 cm	NC
Weight / height	NC	NC	NC	15 kg/99 cm	NC
Growth retardation	+	-	-	-	-
Seizures (age of onset)	(3 months)	-	-	-	-
Intellectual disability	Severe	Moderate	Moderate	Mild	Mild
Early-onset encephalopathy	+	-	-	-	-
Neurological examination					
Abnormal Tone	ND	Axial hypotonia with spastic diplegia	Axial hypotonia with spastic diplegia	-	-
Muscle weakness	ND	+	+	-	-
Movement disorder	ND	Ataxia, tremor	Ataxia, tremor	Ataxia, tremor	Ataxia, tremor
Eye abnormalities	Optic atrophy	Optic atrophy, squint	Optic atrophy, squint	Optic atrophy, squint	Squint
Speech	Delayed	Dysarthria	Dysarthria	Delayed	Delayed
Deafness	-	-	-	-	-
Other clinical observations					
Clinodactyly of 5th fingers	-	-	-	-	-
Other skeletal abnormalities	-	-	-	-	-
Facial dysmorphism	-	-	Long face	-	-
Other	Died at age of 2 years	-	-	EEG normal	-
Neuroimaging					
Brain MRI abnormalities	Diffuse brain volume loss	NA	Diffuse brain volume loss	Cerebellar hypoplasia; posterior atrophy of corpus callosum	NA

**Nomenclature HGVS V2.0 according to mRNA reference sequence NM_006395.2. Nucleotide numbering uses +1 as the A of the ATG translation initiation codon in the reference sequence, with the initiation codon as codon 1.* OA, optic atrophy; CPEO, chronic progressive external ophthalmoplegia; MRI, magnetic resonance imaging; SNHL, sensorineural hearing loss; NA, not analysed; NC, not communicated; ND, not determined

Clinical features of patients with recessive *ATG7* variants.

4. Additional acknowledgement of infrastructure support

The NIHR Newcastle Biomedical Research Centre (BRC) is a partnership between Newcastle Hospitals NHS Foundation Trust and Newcastle University, funded by the National Institute for Health Research (NIHR). This paper presents independent research supported, in part, by the NIHR Newcastle BRC. The views expressed are those of the author(s) and not necessarily those of the NIHR or the UK Department of Health and Social Care.

5. References

1. DePristo MA, Banks E, Poplin R, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat Genet 2011;**43**:491-8.
2. Van der Auwera GA, Carneiro MO, Hartl C, et al. From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. Curr Protoc Bioinformatics 2013;**43**:11.10.1-11.10.33.
3. McKenna A, Hanna M, Banks E, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res 2010;**20**:1297-303.
4. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 2009;**25**:1754-60.
5. Picard. Broad Institute.
6. McLaren W, Gil L, Hunt SE, et al. The Ensembl Variant Effect Predictor. Genome Biol 2016;**17**:122.
7. Gonzalez M, Falk MJ, Gai X, Postrel R, Schule R, Zuchner S. Innovative genomic collaboration using the GENESIS (GEM.app) platform. Hum Mutat 2015;**36**:950-6.
8. Monies D, Abouelhoda M, AlSayed M, et al. The landscape of genetic diseases in Saudi Arabia based on the first 1000 diagnostic panels and exomes. Hum Genet 2017;**136**:921-39.
9. Chen S, Zhou Y, Chen Y, Gu J. fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics 2018;**34**:i884-i90.
10. Dobin A, Davis CA, Schlesinger F, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 2013;**29**:15-21.
11. Wang L, Wang S, Li W. RSeQC: quality control of RNA-seq experiments. Bioinformatics 2012;**28**:2184-5.
12. Amemiya HM, Kundaje A, Boyle AP. The ENCODE Blacklist: Identification of Problematic Regions of the Genome. Sci Rep 2019;**9**:9354.
13. DeLuca DS, Levin JZ, Sivachenko A, et al. RNA-SeQC: RNA-seq metrics for quality control and process optimization. Bioinformatics 2012;**28**:1530-2.
14. Brechtmann F, Mertes C, Matuseviciute A, et al. OUTRIDER: A Statistical Method for Detecting Aberrantly Expressed Genes in RNA Sequencing Data. Am J Hum Genet 2018;**103**:907-17.
15. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 2014;**15**:550.
16. Raudvere U, Kolberg L, Kuzmin I, et al. g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). Nucleic Acids Res 2019;**47**:W191-W8.
17. Luhr M, Szalai P, Saetre F, Gerner L, Seglen PO, Engedal N. A Simple Cargo Sequestration Assay for Quantitative Measurement of Nonselective Autophagy in Cultured Cells. Methods Enzymol 2017;**587**:351-64.
18. Luhr M, Szalai P, Engedal N. The Lactate Dehydrogenase Sequestration Assay - A Simple and Reliable Method to Determine Bulk Autophagic Sequestration Activity in Mammalian Cells. J Vis Exp 2018;**137**:57971.